

Conference Abstract

Answering Research Questions Using Semantic Traits

Paula Mabee [‡]

[‡] University of South Dakota, Department of Biology, Vermillion, SD, United States of America

Corresponding author: Paula Mabee (pmabee@usd.edu)

Received: 21 Jun 2019 | Published: 02 Jul 2019

Citation: Mabee P (2019) Answering Research Questions Using Semantic Traits. Biodiversity Information Science and Standards 3: e37567. <https://doi.org/10.3897/biss.3.37567>

Abstract

Phenoscape has developed methods to render phylogenetic characters from the systematic literature machine computable and interoperable with genetic data from model organisms, by annotating them with taxon, anatomy, quality, other ontologies. Moving these trait data into a semantic framework enables their integration with other data types and provides the potential for powerful new computational tools to aid discovery. For example, trait similarities can be quantified, assessed against phylogenetic trees to determine whether they are based in homology or homoplasy, and linked back to candidate genes. Another example is the ability to automatically construct a matrix on the fly, for a user-selected set of traits and taxa. Guidelines for consistent representation of characters have been developed through manual annotation of over 20,000 systematic characters. These represent a limited number of design patterns that are applicable to traits from any source. The level of detail to which characters are annotated will influence how they may be used in research; a minimal approach will still enable basic trait aggregation. Manual curation effort is substantial and does not scale well to biodiversity traits. Natural Language Processing (NLP) methods, using our newly developed Gold Standard for semantic traits, can accelerate annotation from published text, particularly with well-provisioned ontologies. Efforts to establish new trait databases might profitably explore machine learning for morphological discrimination and semantic annotation from digitized images for a high-throughput approach.

Keywords

ontologies, systematics, characters, traits, Phenoscape

Presenting author

Paula Mabee

Presented at

Biodiversity_Next 2019

Acknowledgements

The Phenoscape team, including T.J. Vision, H. Lapp, J. Balhoff, and W. Dahdul contributed significantly to the research described.

Grant title

NSF 1062542; Collaborative research: ABI Development: Ontology-enabled reasoning across phenotypes from evolution and model organisms

Hosting institution

Department of Biology, University of South Dakota, Vermillion, SD 67069 USA

Ethics and security

None

Author contributions

Paula Mabee authored the abstract

Conflicts of interest

none